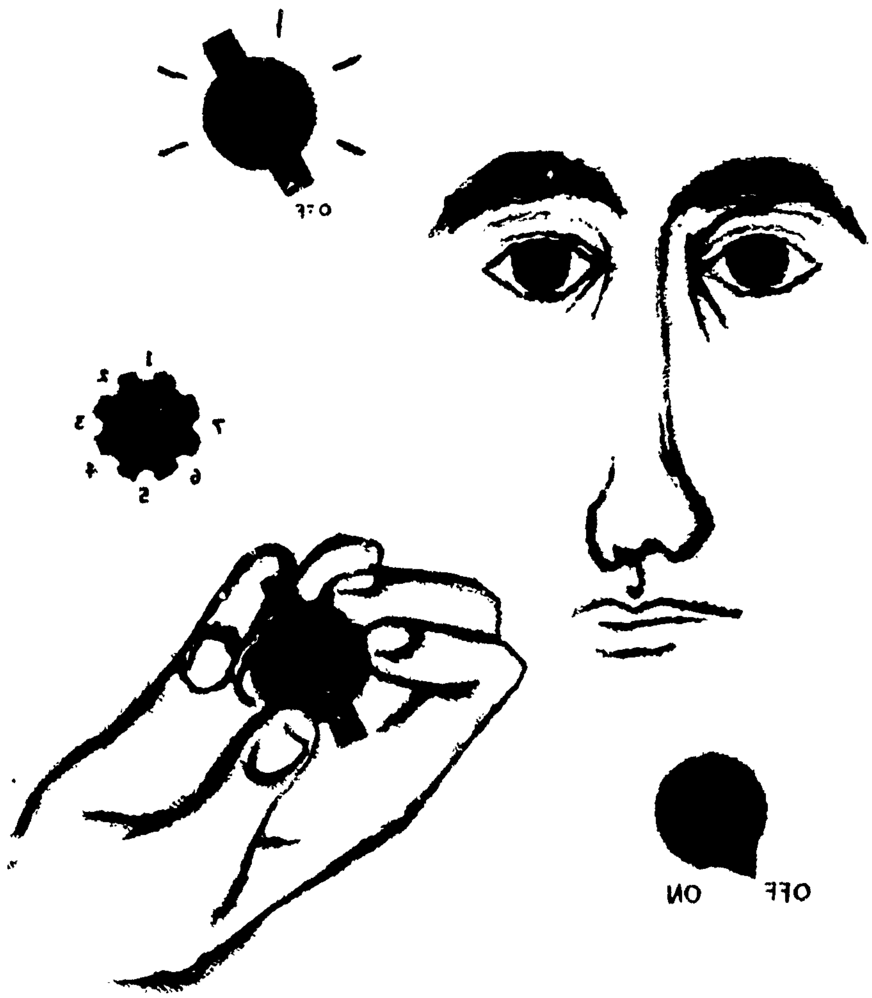
AD 607164

an index of
electronic
equipment
operability

REPORT O
DEVELOPMEN



AIR-C-43-1/62-FR

An Index of Electronic Equipment Operability

REPORT OF DEVELOPMENT

Dan Payne
James W. Altman

American Institute for Research
Pittsburgh, Pennsylvania

31 January 1962

Contract No. DA-36-039-SC-80555
Signal Corps Technical Requirements EWD
Spec. No. 3-60 dated 15 February 1960

FOREWORD

This report is concerned with the development of the Index of Electronic Equipment Operability. It is one of five related documents. The Instruction Manual, Data Store, and Evaluation Booklet are required for evaluating equipment. The Sample Equipment Evaluations report contains detailed evaluations of four equipments, including recommendations. This work was performed under Contract No. DA-36-039-SC-80555 for the Electronic Warfare Department, United States Army Electronic Proving Ground, Ft. Huachuca, Arizona. Mr. James J. Edwards and Walter Bonham served as Technical Representatives of the Contracting Officer, and provided continuing support during the conduct of the study. Mr. Paul Lamb, Electronic Warfare Department, and Mr. Jeff Abraham, Signal Communications Department, were of considerable assistance, serving as evaluators during the tryout of the Index. The authors are also indebted to numerous personnel of the American Institute for Research, but especially to Mr. Manus R. Munger for his critical review and general contribution to the preparation of the Instruction Manual, and to Mr. Robert W. Smith and Mrs. Sara J. Munger, members of the project staff, for their over-all support and contribution throughout all phases of the effort.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

PURPOSE

The Signal Corps has long been aware of the need for evaluation of its equipment. The evaluation of hardware has become integral to the development process, and the recent past has shown the development and implementation of a number of general procedures which essentially evaluate the human component of systems. However, electronic equipment has become so complex in recent years that general, informal evaluation procedures are no longer feasible. The nature, number, and inter-relationships of factors prevent adequate overall evaluations, and the tolerances of existing and future equipments demand quantitative information which such procedures cannot provide. Recently, a quantitative procedure for the evaluation of electronic equipment maintainability (Munger & Willis, 1959) was developed and implemented. The purpose of the present study is to provide a procedure for the early, quantitative evaluation of electronic equipment <u>operability</u>.

<u>Previous Efforts</u>

In recent years there have been numerous attempts to assess the human element of systems. Handbooks and guides to aid design engineers have been prepared in considerable numbers (Baker & Grether, 1954; Ely, Thompson, & Orlansky, 1956; Van Cott & Altman, 1956; Folley & Altman, 1956; Altman, et al, 1961). Similarly, there has been a considerable effort to develop human factors checklists for use during evaluation (Krumm & Kirchner, 1956; Berkun & Van Cott, 1956; Van Cott, 1956; Fitzpatrick, 1955; etc.). While these procedures fulfilled a known need, they are generally inadequate for purposes of evaluation. Although the information presented is often based upon experimental comparisons, there is no way of knowing the consequences on actual performance.

Also, as pointed out by Shapero & Bates (1959), "it has been difficult to integrate the human elements with the rest of the weapon system." Shapero & Bates develop a "system analysis and integration model" to overcome this difficulty. In a sense, they achieve their purpose. There model does integrate the human element with the remainder of a system, but in a qualitative manner.

That is, although they can demonstrate the interaction of the human element with all aspects of the system, their scheme does not provide information about the consequences of the interaction. The chief difficulty lies in the lack of comparable data about human and other system elements. The performance of other elements is generally well known, and quantified. The performance of the human element is generally neither.

In the past few years, there have been several attempts at quantifying the human element. These efforts are typified by the work of Williams (19— Kaufman, Oehrlein, & Kaufman (1961), and Siegel & Wolf (1961). While notab— in concept, these procedures are either too gross, or require information — is generally not available at the time of evaluation.

Williams has proposed a human reliability evaluation procedure based or equipment reliability assessment procedures. However, the reliability f—ur employed in his procedure are estimates to be made by the evaluator. It is thus doubtful that two independent evaluations of the same equipment would — similar. Kaufman, Oehrlein, & Kaufman have based their procedure on easily available information. Here, human reliability is related to such factors — volume, cost, and weight of the equipment to be operated. The assumption t— these factors accurately reflect design sophistication seems questionable. And the further assumption that design sophistication is directly related to human reliability seems untenable if field operation is the criteria.

The computer simulation approach proposed by Siegel & Wolf is a unique attempt to integrate a notion of performance consequences with other system considerations. However, their concern is with the determination of "operat— overloading" based upon estimates of human performance time and error. That is, if performance time and/or errors are in excess of tolerances, then the operator is overloaded. For this purpose, the model is appropriate. It con— tains some useful notions for purposes of evaluation, but in itself is not appropriate as a general evaluation technique.

<u>Problem</u>

In view of the general purpose of the study, and the experiences of oth— the primary problem here was to develop an evaluation procedure which provi— quantitative information related specifically to operator performance. It —

planned that the procedure should be applicable during, or prior to, acceptance testing. Objectives for the procedure were to:

1. Predict the time and reliability (accuracy) of operator performance.
2. Identify specific design features which degrade operator performance.
3. Provide general guidance concerning selection and training of operators for evaluated equipments.

<u>Requirements for an Operability Evaluation Procedure</u>

Criteria which guided development of the Index were:

1. <u>Meaningfulness</u>. Results should be in terms such as speed and accuracy of performance which are directly meaningful rather than in indirect measures that would require considerable interpretation.
2. <u>Specificity</u>. The specific design features and aspects of performance contributing to operational complexity should be made explicit in the evaluation process, resulting in a diagnostic as well as an over-all evaluation tool.
3. <u>Objectivity</u>. Sufficient guidance should be provided to permit exactly the same evaluation results by independent evaluators, within the limits imposed by their irreducible observational and judgmental differences. On the other hand, the evaluator should be permitted to note any instances in which he feels the formal evaluation procedure is incomplete or would be misleading without specific interpretation.
4. <u>Comprehensiveness</u>. Although there is no way to guarantee that every important factor will be assessed, every factor of known importance should be included in the procedure.
5. <u>Ease of use</u>. Every effort should be expended to make application of the evaluation procedure as simple and straightforward as possible through the preparation of guidance materials and forms.

It was felt that if the above criteria were considered throughout the developmental process, there would be maximum likelihood that the resulting procedure would be both reliable and valid. By reliability in this context is meant the extent to which results of independent evaluations for the same item of equipment will be similar. By validity is meant the degree to which evaluation results will accurately predict actual operator performance. Although there was only very limited opportunity within the scope of this project to study reliability or validity through empirical tryout, preliminary results are sufficiently promising to suggest further study and refinement.

CONCEPTUAL APPROACH

The Problem

The central problem in the achievement of the above objectives was the development of a conceptual framework which would enable the use of existing experimental data. This problem had two major implications:

1. Since the bulk of human engineering experimentation is concerned with the details of hardware design, any basic framework used in development of the evaluation procedure must include a breakdown of general hardware categories and characteristics.

2. Although the ultimate goal is to predict mission performance, there are no human performance data available at this level, and little if any usable data even at the mission phase or task level. Consequently, the finest unit for which reasonable performance data can be established is the individual step, act, or behavior.

Aspects of Behavior

Even for the individual task step or behavior, the available performance data are generally not appropriate. In order to organize the existing data in any useful fashion, it seemed necessary to consider the following aspects of each behavior separately:

1. Reception of information relevant to the behavior.
2. Internal processing.
3. Responding.

Changes in any of these aspects alter the nature of the behavior unit, and should be reflected in predictions relating to its performance.

These aspects generally fit the standard STIMULUS-ORGANISM-RESPONSE paradigm. A brief attempt was made to define discrete units of behavior in terms of these aspects. But, while a good deal is known about vision, audition,

perception, decision making, and psycho-motor activity, the knowledge as well as the theory, is not yet sufficient to handle practical problems at this level. However, these aspects, translated into hardware terms, seemed quite practical.

Generally, man is associated with a machine input via a control, and with the output via a display of some sort. That is, the man's input (stimulus) is the machine's output, and the man's output (response) is the machine's input. It seemed, therefore, that a careful study of the sources of machine outputs would provide the information concerning the range of stimuli with which men would be expected to cope. Similarly, a study of machine inputs, essentially controls, would identify a majority of the characteristics of man's response. Thus, the SOR concept was expressed in terms of the <u>source</u> of the stimulus, <u>gross</u> mediating processes, and the <u>mode</u> and <u>media</u> of the operator response.

Ultimately, the framework for performance analysis involved four levels of classification:

1. <u>Aspects of behavior</u> which refer to categories of inputs, mediating processes, and outputs.
2. <u>Components</u> which refer to a specific category of an aspect, e.g., joystick is a component of the output aspect.
3. <u>Parameters</u> which refer to the relevant characteristics of components, e.g., stick length is a parameter of the component joystick.
4. <u>Dimensions</u> which refer to specific values of the relevant parameters, e.g., six inches is a dimension of the parameter stick length.

Required Information

The information required by this conceptual approach is of several varieties. The evaluator must have detailed equipment information that is relevant to operation. This is generally restricted to design details of controls and displays and their spatial and functional relationships.

Detailed information is also required about the use of these controls and displays. Essentially, this is operating information contained in a task analysis or detailed operating manual.

Of primary importance here is the need for performance data relevant to operation. That is, there must be a convenient store of information that contains performance data for any control or display that may be encountered. In addition, there must be guidance materials to insure consistency of the evaluation process and its results, as well as aiding the interpretations and use of the results. Figure 1 presents this gross conceptualization graphically.

Limitations

The major limitation inherent in this approach is that the consequences of specific components and parameters _in interaction_ are unknown. The determination of these effects, at the level of detail required here, is currently beyond the state-of-the-art. It is assumed that interaction effects will tend to balance out so that results of evaluation will not be consistently in error.

A second potential limitation is due to the reliance upon experimental data. Insofar as possible, available experimental data were used in formulating guidance for performance estimates. Thus the final procedure, and its results, can only be as good as the data upon which it is based. The major work is yet to be done in establishing reliable, general standards of human performance.

Additional limitations of the Index are inherent in the statement of assumptions below.

Assumptions

To assure that the Index achieves its stated purposes, the following conditions must be met in applying the Index:

1. Available equipment and task information must accurately describe the design and operating characteristics of the equipment to be evaluated. Any change in the design of the equipment or the allocation of operator responsibilities will alter the detailed evaluation results and may significantly alter the interpretation of the results.

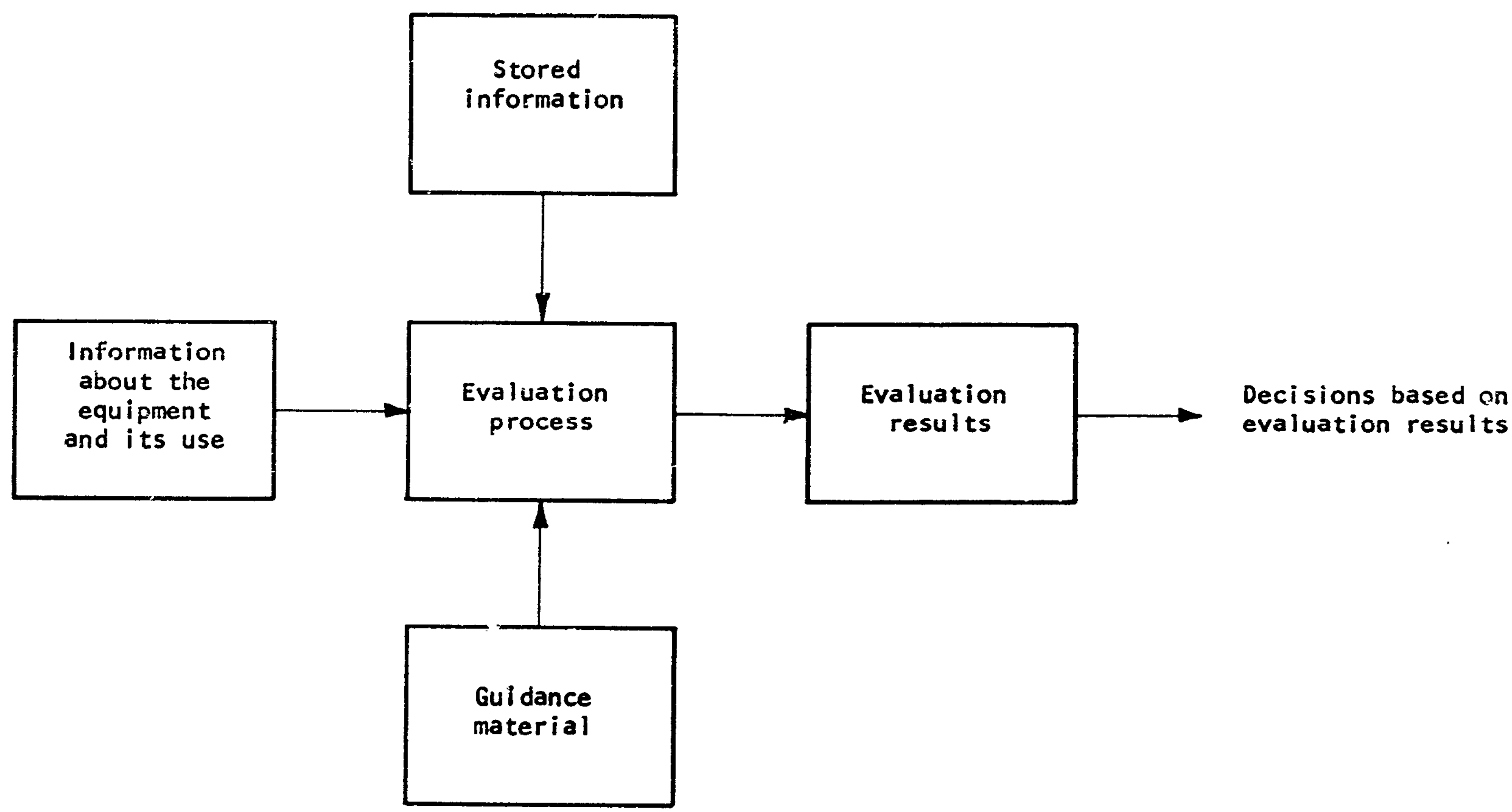


Figure 1. Gross Schematic of the Operability Evaluation Problem

2. The Index is intended for prediction of performance by relatively unselected personnel who have received only nominal training. In most cases, rigid selection criteria or intensive training will result in operator performance that is faster and more reliable than performance predicted by the Index.

3. The Index should be applied by a professional human factors engineer or other personnel qualified to evaluate man-machine interactions.

DEVELOPMENT OF THE INDEX

Summary

Measurable performance related equipment components, applicable to any operating behavior, were identified and categorized in accordance with the general framework. General and experimental information pertaining to these characteristics and the factors, or parameters, affecting performance were abstracted from the literature. This abstracted data was related to the categories of equipment components. A general correction factor was computed and applied to all data to compensate for the laboratory conditions under which they were generated. The resulting corrected data was integrated and organized into a store of data for ready access. Procedures for evaluation based upon this data and equipment and operating information were developed. Scoring procedures were developed in detail, and guidelines to the interpretations of evaluation results for the specified purposes were presented. Each major effort is covered in more detail below.

Components of Behavior

Identification of Components

General performance related equipment components, or behavior components, were identified. Various types of equipment and equipment manuals were surveyed. During the survey, all controls nd displays observed were noted. In addition, the "thinking" or mediating process required of operators was inferred and noted. The resulting lists were somewhat lengthy, due to numerous variations of a few basic components. The categories of components selected for inclusion in the Index consisted, generally, of these basic, unique components. In some cases, however, subsequent experimental findings led to a further breakdown of some components. For example, the general component "scales" was subsequently broken down into types of scales. These categories were then related to input, mediating process, or output aspects in accordance with the basic framework. The result of this process appears in Table I below.

Table I. List of Input, Mediating Process, and Output Components

Inputs	Mediating Process	Outputs
Circular Scales	Identification/Recognition	Cable Connections
Counters	Manipulation	Cranks
Labeling		Disconnecting
Lights		Joysticks
Linear Scales		Knobs
Non-Speech		Levers
Scopes		Object Positioning
Semi-Circular Scales		Pushbuttons
Speech		Rotary Selectors
		Speech
		Toggle Switches
		Writing

These components, and their variations, account for most of the important sources of information to the Operator (input), his treatment of this information (mediating process), and the modes of his responses (output).

Parameters Affecting Component Performance

For each component, the associated parameters or factors which affected performance were identified. The approach here was both rational and empirical. The rational approach consisted of a careful consideration of each component. The attempt was to identify all possible factors that might affect the use of a component. These parameters were noted and checked against the empirical approach results.

The empirical approach consisted of surveying the experimental literature to identify the dependent variables which had been studied. With 15-20 years of experimentation as a source, it was felt that most significant components and factors had already been identified and studied. Thus, the literature was reviewed and all studies concerned with a given component abstracted. The resulting abstracts were summarized so that all the parameters studied and their consequences for performance could be observed.

The resulting list of parameters was then compared with the parameters identified by rational analysis. The purpose of the comparison was to insure that all parameters of significance had been studied and the data abstracted. Some parameters had apparently never been studied, and other were impossible to isolate or measure without elaborate equipment. Where there was no data on a parameter, its effect was judged; where the parameter could not be readily measured, gross notions as to its dimensions were identified which could be readily judged by an evaluator.

As an example of the kinds of parameters identified, Table II presents the parameters affecting performance on the component "lights."

Table II. Parameters Affecting Performance on Lights

LIGHTS	Size Brightness Type/Function Number Presentation

These parameters are believed to be the most important ones in terms of their effect on performance, and they can be easily identified.

An attempt was made to incorporate other factors, such as situationa motivational, personality factors, etc., into the Index. This effort was abandoned, however, when it became clear that existing data was quite contradictory and insufficient for Index purposes.

Performance Data

Data Abstracts

Performance data related to the components and parameters identified were abstracted from the experimental literature. Over the course of the study, several thousand research reports were surveyed. Of these, severa hundred were selected for careful consideration. Reports meeting the following requirements were finally abstracted.

1. Experimental in nature.
2. Specific to type(s) of control or display, or generalizable

to input, mediating process, or output aspect of behavior.

3. Raw or grouped data presented with the analysis, rather than simple reports of conclusions.
4. Emphasis on time and/or error measures, or measures translatable into these terms.
5. Well-defined dimensions of controls and displays.
6. Explicit statement of experimental method and conditions.

A total of 164 research reports meeting these requirements were abstracted. References to these reports appear as an Appendix.

Two examples of the kind of data available in the literature, and the way it was extracted, are presented in Figure 2.

Figure 2. Sample of Abstracted Data

Example 1

If a rotary knob is used for making settings on a linear scale, is the control friction approximately 100 grams (at periphery of knob)?

If this is not followed:

Operation or setting times will be increased (if greater friction is used).

Mean Total Time (seconds)

Friction	10/16" travel	50/16" travel
100 grams	2.24	2.92
400 grams	2.38	3.90*
700 grams	2.45	4.45*
1,000 grams	2.37	4.81*
1,300 grams	2.58*	5.10*

*Significant differences beyond 1% level from underlined figure in column.

(Jenkins, L. J. J. Appl. Psychol., 1950, 34)

Example 2

Ss: 20 college students with normal Snellen Acuity and no obvious visual defects.

Task:

Photographs of circular moving pointer dials were presented in a slid mirror tachistoscope. Viewing distance was 28"; viewing angle was varied from 90° to 25°. Two types of white on black dials were used: (dial diameter was not given).

1. 600 unit dial-- graduation marks every 10 units; 50 and 100 marks heavier and longer; 100 unit marks also numbered.
2. 400 unit dial--graduation marks every 10 units; every 40 unit markers heavier and longer and numbered.

Brightness of white markings was 7 ft. lamberts.

Ss were given 10 practice trials. Each S was given 40 test tials, 20 on each dial type, 4 at each of 10 viewing angles, and 5 in each dial quad rant. For each dial, half the settings were near a graduation mark and half were near a mid-mark position. S controlled exposure time; E record time and the reading given by S. Instructions were for S to "read the dial to the nearest 5 units as accurately and quickly as possible."

Per Cent Reading Errors of 5 Units or More at Each Viewing Angle
(No. Ss = 20. No. readings = 4 per S at each viewing angle)

	Viewing angle (Both types dial combined)									
	90°	80°	70°	60°	50°	45°	40°	35°	30°	25°
% Readings in Error by 5 Units or More	14.0	12.5	15.0	21.0	17.5	16.5	22.5	23.5	20.5	22.5

(Data extrapolated from graph.)

Reading time showed no systematic change associated with viewing angl with either dial.

(Cohen, Vanderplas, & White. J. Appl. Psychol., 1953, 37)

Ideally, a comprehensive collection of such data could be treated to yield performance data for each equipment characteristic, considering all relevant factors which might influence performance. In reality, however, abstracted from the literature is markedly lacking in consistency with res to the factors investigated, kinds of measure used, and experimental rigor Human engineering studies, which yield the most relevant information, are minimally generalizable. Most of these studies were conducted to answer c specific questions, and there is a marked lack of any theoretical framewor within which such studies were conducted. Thus, while a great amount of

relevant experimental data exists, it was difficult to bring the data, in its original form, to bear directly on the general problem under study.

There was, then, the significant problem of reducing and integrating the available data into a form compatible with an evaluation procedure. In approaching this problem, the performance measures used by the various investigators were transformed into consistent time and error terms. That is, if at all possible, dependent measures were expressed as per cent of trials in error, and time required per trial. In only a few cases was this not possible. This translation into consistent measures tended to reduce the complexity of the massed data.

Reduction of Abstracted Data

Reduction of the abstracted data was required to reduce the mass of apparently unrelated information. Tables were prepared which summarized, for each component, all available abstract information. These tables presented:

1. A list of possible parameters for each component.
2. The dimensions, or specific values, of each parameter studied and performance measures related to each dimension.
3. The experimental conditions of each study and the performance measures related to each condition.
4. The number and kinds of subjects, number of practice and test trials, etc., and related performance measures.

Both discrepancies in data and the absence of data could be easily noted from the tables. A majority of the discrepancies were found to be due to gross differences in the mediating processes required. Imposing a consideration of these processes on the data eliminated most of the major differences. Those few instances where this was not the case were reconciled by the judgment of the project staff. The absence of data resulted in a specific search of the literature to fill the gaps in the summary tables. Where this search failed, the data were generated by extrapolation or interpolation from related studies, or, as a last resort, by expert judgment.

The major reduction of data was accomplished by grouping together some dimensions of the studied parameters. It was obvious that the parameters could not, and should not, be presented at the detailed level at which they were studied. The decisions for grouping the dimensions into a workable number were somewhat arbitrary. Every effort was made to base grouping on the statistical significance of differences found between dimensions. Where this was not possible, the criterion of meaningfulness of dimensions took precedence. An example of how this dimensional grouping was accomplished is presented in Table III below. The data refers to control knobs. The parameter of concern in this example is "size."

Table III. Example of Dimensional Grouping

Average Knob Turning Time (in Seconds) Under Varied Shaft Friction

Original Data

Knob Diameter	Moderate Friction	Heavy Friction
1/2 inch	1.649	2.170
3/4 inch	1.553	1.802
1 inch	1.318	1.585
1 1/4 inch	1.237	1.498
1 1/2 inch	1.262	1.368
1 3/4 inch	1.213	1.328
2 inch	1.211	1.264
2 1/4 inch	1.208	1.281
2 1/2 inch	1.256	1.317
2 3/4 inch	1.245	1.430
3 inch	1.292	1.419
3 1/4 inch	1.275	1.394

Grouped Data

		Widest Difference Between Original Means			Widest Diffe Between Origina	
	Mean Operation Time	Actual	As % of Mean Operation Time	Mean Operation Time	Actual	As % of Operatic
Less than 1"	1.601	0.096	0.06	1.986	0.268	0.1
1" to less than 2"	1.257	0.105	0.08	1.445	0.257	0.
2" to less than 3"	1.230	0.048	0.04	1.323	0.166	0.
3" or more	1.283	0.017	0.01	1.406	0.025	0.

From this table it is evident that little loss of data occurred by grouping in this manner. The range of operation time included in the groups is a very small percentage of the mean operation time for a given size of knob. Yet, differences in time between knob sizes are apparent. In those few cases where several studies showed divergent grouping tendencies, final groupings were decided arbitrarily.

The result of this process was a first approximation store of data for use during evaluation. However, the data at this point was still expressed in terms of time per trial, and per cent of trials in error. In order to isolate the contribution of a given dimension to time and error, further integration of the data was required.

Data Integration

The most frequent case among the studies abstracted was where two or more parameters were varied simultaneously. It was necessary to determine the general, but independent effects of each dimension of every parameter upon both total time and error. The integration procedure was as follows.

The magnitude of the measures obtained in the abstracted studies is dependent upon known dimensions of parameters being studied, and all other factors and conditions, whether controlled or not. It is reasonable to assume, however, that within a given study, these other factors remain fairly constant. Therefore, differences in obtained measures may be attributed to known variations of dimensions of parameters being studied. A change in conditions may be expected to alter the magnitude of the measures, but the differences attributable to known variations, relative to the magnitude, would be expected to remain the same, except as a result of their interaction with the new condition.

Time Estimates. Time data from a single study on "rotary controls" is presented in Table IV to serve as an example of how time consequences were determined.

Since, as stated above, the magnitude of the measures is dependent on a host of factors, most of which are unknown, there is a minimal concern with the actual numbers in Table IV. Differences in numbers attributable

Table IV. Sample Data on Rotary Controls

	Size		Placement			Time In Seconds
	<1"	2-3"	Front	Top	Side	
	X		X			2.78
	X			X		3.55
	X				X	3.27
		X	X			3.08
		X		X		2.96
		X			X	3.94
Means	3.2	3.33	2.93	3.25	3.6	3.26

to the dimensions studied are determined by taking differences between means. That is, with regard to size, the difference between the means for <1" and 2-3" controls is .13 seconds. In this case, we say that the consequence of 2-3" controls, rather than <1", over placement and all other unknown factors and conditions, is .13 seconds. The consequence of <1" controls can only be interpreted as zero, since in this example it is the optimum dimension. With regard to placement, the Front is optimum here, and its consequence is zero. The performance consequence of Top is .32, and Side is .67 seconds. Since it is most reasonable to establish consequence for deviations from an optimum, there is no concern with differences between Top and Side placement.

Optimum levels of performance for each component were determined. That is, a "base time" was established for each component which assumed all parameters were optimum. This base time was determined by searching the abstracts and finding optimum conditions for the component under study. Given a base time, the consequences of non-optimum dimensions of parameters were considered as time added to the base.

Such consequences and base times were determined for all components. Where replications of studies occurred, mean consequences were established. Where a given factor was studied under significantly different conditions, the abstracts of the studies were examined. Generally, it was possible to justify and adjust consequences based on obvious factors such as subjects or practice trials, etc. In the few cases where consequences appeared irreconcilable, adjustments were made based on judgment.

Error Estimates. Time measures obtained in this manner seemed to be reasonable estimates of performance time under actual operational conditions. The error estimates, however, were considered gross over-estimates of operational errors, because of artificial inflation of error counts. This inflation occurs because error rates are normally relatively low in operational task performance. In order to have measurable error without running a very large number of trials, experimenters inflate errors by making tasks unusually difficult, or counting potential, or near, errors. The derivation of some operational meaning from this data, relevant to the evaluation problem required a more devious approach.

The most meaningful notion of operator accuracy for evaluation purposes is that of operator reliability. But this notion of operator reliability should apply to the components of behavior. That is, reliability measures should be available for each of the inputs, mediating processes, and outputs. In order to achieve this, the reliability contribution of each dimension of every relevant parameter must be known. Ideally, the reliability contribution at this level should be determined empirically. However, this would require a long term, extensive effort which was far beyond the scope of the study. The interim solution to this problem consisted of scaling the grossly inflated laboratory error counts against available estimates of over-all field reliability.

The over-all estimates of field reliability were obtained from previous studies (Miller, et al, 1957, and Craig, et al, 1957). Over a variety of equipments and missions, the range of operator reliability estimates was between 85 and 90 per cent. Conversely, it may be said that 10 to 15 per cent of the time, operator error will fail or seriously degrade mission effectiveness. No field studies have been conducted which provide reliability estimates for individual task steps, behaviors, or behavior components. The best estimate at this level seemed to be a "mean mission step unreliability" figure. Unreliability was chosen for computational convenience only.

This estimate was taken by determining the mean number of steps in a mission, and dividing by a mean mission unreliability estimate of .13. The mean number of steps was determined by counting the required steps for 26 different equipments. The mean number was near 50. The mean mission step unreliability obtained in this way was .0026. In different words,

the best estimate seemed to be that 26 times in ten thousand, an operator error on a given <u>step</u> of operation would fail or seriously degrade mission effectiveness.

This mean mission step unreliability was then compared with an estimate of mean unreliability per experimental trial. This was determined from the data available from the abstracts. Over the data available, it was found that mean unreliability per trial was .31935. Thus, there were two estimates of mean step unreliability--one based on estimates of actual field operation, one based on laboratory experimentation. Assuming that experimental trials are roughly equivalent to individual steps of operation, the ratio of these means is a reasonable conversion factor for laboratory results. That is, correcting all the experimental results by a factor of $\frac{.0026}{.31935}$, or .008145, compensated for the laboratory conditions, and rendered the data more compatible with field operation. The corrected unreliability figures were then converted to conventional reliability scores.

<u>Attributing effects to behavioral components</u>. These reliability scores, based upon steps of operation and experimental trials, were considered attributable to individual components of behavior. The justification for this is that, in an experiment involving an input, every effort is typically made to reduce error or unreliability due to mediating and output aspects of the behavior to an absolute minimum. Expressed in other terms, the reliability of a behavior is dependent upon the reliabilities of the aspects of behavior. In experimental studies involving one aspect, the reliabilities of the other aspects are made to approximate unity. When time is the dependent measure for an aspect under study, the time attributable to the other aspects is held to a minimum. Thus, attributing time and/or reliability, as corrected by the above factor, to a single component seems reasonable.

Organization of the Data Store

Organization of the data into a convenient, accessible form was possible with the treated data. The result of this organization was the Data Store of the Index. Figure 3 presents the data as it was finally presented in the Data Store. Individual card(s) were prepared for each component. On this card, the parameters relevant to the component were presented. The dimensions associated with each of the parameters were listed with the associated

data. The base time for each component appears at the top of the time column. This base, as mentioned previously, serves as an absolute minimum time for behaving with the component. This time will hold only if all the parameters listed are of optimum dimension, e.g., add no time. Other dimensions will add time to the base. Reliability estimates are presented for each parameter dimension.

Figure 3. Sample Data Store Card

JOYSTICK
(May move in many planes)

BASE TIME = 1.93

Time added	Reliability	
		1. Stick length
1.50	.9963	a. 6-9"
0	.9967	b. 12-18"
1.50	.9963	c. 21-27"
		2. Extent of stick movement (Extent of movement from one extreme to the other in a single plane.)
0	.9981	a. 5-20^{o}
.20	.9975	b. 30-40^{o}
.50	.9960	c. 40-60^{o}
		3. Control resistance
0	.9999	a. 5-10 lbs.
.50	.9992	b. 10-30 lbs.
		4. Support of operating member
0	.9990	a. Present
1.00	.9950	b. Absent
		5. Time delay (Time lap between movement of control and movement of display.)
0	.9967	a. .3 sec.
.50	.9963	b. .6-1.5 sec.
3.00	.9957	c. 3.0 sec.

Information Required for Evaluation

Two general types of information concerning the equipment to be evaluated must be obtained before the Index can be successfully applied.

Equipment Information

Data concerning the equipment should include detailed information about the controls and displays. If prototype, pre-prototype, or mock-up equipment

is available for evaluation, this information can be obtained directly. If an evaluation is to be conducted prior to mock-up, then the level of detail required may present a problem. The Data Store described above indicates the nature of the equipment information required.

Operating Information

This information is concerned with what an operator must do with the controls and displays. Most of the information will be contained in a good task analysis, or in detailed operating manuals, or may be supplied by an expert on operation of the equipment. In all cases, however, some of the information must be inferred. While the input and output of operation are almost always easily specified, the mediating processes involved in operation must be inferred. At the level of concern here, however, the inferences are not difficult to make.

Guidance Materials and the Evaluation Process

Given the Index Data Store and the information required concerning the equipment and its operation, the evaluation process becomes essentially that of matching the information with the data. The guidance materials developed are detailed instructions to guide this process. In addition, instructions are presented for scoring the Index, and interpreting the results.

The application of the Index requires the completion of six major steps or processes. These steps are listed briefly below. Detailed instructions for each step are contained in the Index Instruction Manual.

1. Organize Equipment and Operating Information. Data obtained from task analyses and other sources must be analyzed into behavioral steps and sequenced by mission phases of operation.
2. Collect Evaluation Data. This step includes the identification of relevant components, parameters, and dimensions for each step, matching these values with the data in the Data Store, and entering the appropriate values on an Evaluation Sheet.

3. <u>Score Evaluation Sheet</u>. Step scores are computed for each aspect of behavior and across aspects for total step scores by adding together the relevant time entries and multiplying together the reliability estimates. These totals are entered on the Evaluation Sheet.

4. <u>Summarize Results by Mission and Phase</u>. Total values for each phase of a mission and for the total mission are computed from the data on the Evaluation Sheet, similar to the method for obtaining step totals. The results of this summary are entered on the Mission and Phase Summary sheet.

5. <u>Summarize Results by Component</u>. Total values for each component of the input, mediating process, and output aspects of behavior are computed across the steps of each phase of the mission. The values are entered on the Component Summary Form.

6. <u>Derive Recommendations</u>. Based on the summarized results of the evaluation listed above, recommendations may be developed in the following three areas:

   a. <u>Redesign</u>. Redesign recommendations are based on consideration of total component scores on the Component Summary Form and selection of alternate dimensions from the information contained in the Data Store to improve potential operator performance.

   b. <u>Training</u>. Training recommendations will be based on analysis of the Component Summary Form and will identify aspects of performance that should be given special attention in the training of operators.

   c. <u>Selection</u>. Selection recommendations will be based on identification of aspects of behavior which contribute significantly to total mission scores on the Mission and Phase Summary Form. These aspects may then be related to general selection requirements for operators emphasizing these aspects.

Summary Description

A graphical summary of the basic evaluation process involved in the Index is presented in Figure 4. Essentially, the individual steps of operation are analyzed in their component parts. Scores for these components are determined with the aid of the Data Store. The component scores, and between step time scores, are then combined into step scores. The step scores can then be combined in various ways to yield total aspect, phase, and mission scores. Total scores for specific components are taken from the general component scores. This array of quantitative information of different levels can then be used to guide decisions and recommendations concerning the equipment evaluated.

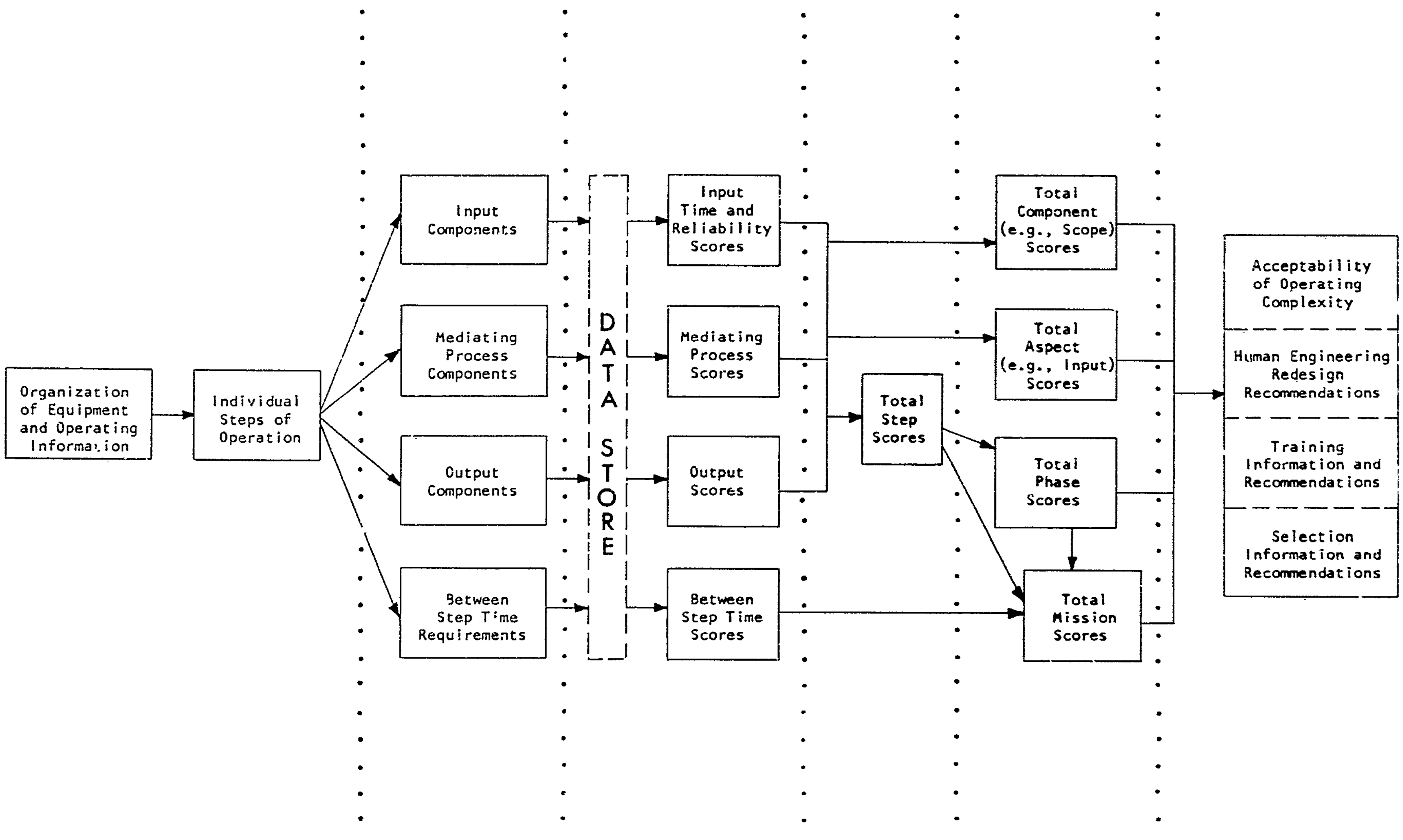


Figure 4. Graphical Summary of the Basic Evaluation Process

INDEX TRYOUT AND RESULTS

During September 1961, a field tryout of the Index was conducted. The tryout consisted of evaluation and reporting on four equipments. They were

1. AN/GRC-50 Mobile UHF Radio Relay Equipment
2. AN/APS-94 Airborne Radar System
3. M-33 Anti-Aircraft Fire Control System
4. AN/MLQ-8 (XL-2) Electronic Countermeasures Set

The evaluation materials and reports, including recommendations, are contained in the Sample Equipment Evaluations Report.

The field tryout of the Operability Index had four major goals. They were:

1. Evaluate the extent to which the Index can be applied to different equipments (versatility).
2. Determine the consistency of time and reliability scores derived by different evaluators for the same equipment (reliability).
3. Determine the extent to which the Index reflected the known operability of the equipment undergoing evaluation (validity).
4. Determine the effectiveness of the evaluation data in diagnosing problems in the area of equipment design, selection, and training (utility).

Versatility

In order to assess its versatility, the Index was applied to equipment which varied in terms of operating requirements. For example, the AN/GRC-5 is operated by one man, and chiefly involves aligning and adjusting the equipment. The M-33 is a multi-man operation concerned with the manual acquisition of targets. A unique aspect of operating, in terms of the Index, was the antenna erection required with the AN/MLQ-8. The AN/APS-94 was significant in that the control panels were largely miniaturized.

In all cases, the equipment evaluations were accomplished with little or no difficulty. The behavior components, parameters, and dimensions of the Data Store were both inclusive and definitive enough to assure an effective evaluation. Subsequent analysis of the evaluations pointed out certain areas and aspects of the evaluation materials that could be revised to further enhance the versatility of the Index. These revisions were chiefly concerned with improving definitions of terms used in the Index.

As a consequence of the tryout and revisions made, it is believed that the Index is sufficiently versatile to be applicable to a wide, if not exhaustive, range of electronic equipments.

Reliability

The evaluations conducted during the tryout were performed by two members of the project staff and two civilian employees of the U. S. Army Electronic Proving Ground. An engineer from the Electronic Warfare Department evaluated two equipments, a Human Factors Specialist with the Signal Communications Department evaluated one equipment. These two individuals had no familiarity with the Index, except that gained from the Instruction Manual and their own experience during evaluation. Members of the project staff, well trained in the use of the Index, evaluated all four equipments. Table V presents a summary of the evaluation results in terms of total scores for each evaluator for each equipment evaluated.

Table V. Total Index Scores for Individual Evaluators

Equipment	Score	Evaluators A	B	C	D
AN/GRC-50	Time	990.91	1206.11		
	Reliability	.47	.46		
M-33	Time	487.90	194.74	191.80	
	Reliability	.82	.81	.81	
AN/MLQ-8 (XL-2)	Time	174.52	166.22		202.38
	Reliability	.96	.96		.96
AN/APS-94	Time	98.08	73.40	140.95	
	Reliability	.96	.95	.96	

The evaluation of the AN/GRC-50 was considered as a preliminary tryou of the Index. Based on the results of this one evaluation, significant changes in the Index were made. A majority of these changes were procedur in nature, based on concepts developed during the evaluation. Changes in Index materials were also made, but were generally minor in nature. The revisions made, however, were numerous enough, and of such significance that the evaluations of the other equipments were entirely different in nature from the evaluation of the AN/GRC-50. Since time prevented a re-evaluation of this equipment, it does not enter in the detailed analysis of results which follows.

Estimates of Inter-Rater Reliability

Inter-rater reliability estimates were computed over the three evalu-tions selected for detailed analysis. These estimates were based on the rank order of both phase and aspect scores. Reliability estimates were computed separately for both time and reliability scores. The results of this analysis are summarized in Table VI below. More detailed tables relating to each of the entries below appear in Appendix B.

Table VI. Summary of Inter-Rater Reliability Estimates for Phase and Aspect Scores. (Median reliability for three evaluators).

Equipment	Level	Reliability of Time Scores	Reliability of Reliability Scores
All	Phase	.86	.96
M-33	Aspect and Phase	.72	.92
AN/MLQ-8	Aspect and Phase	.86	.94
AN/APS-94	Aspect and Phase	.75	.94

The first entry above considers only total phase scores, without reg to equipment type. The remaining entries are based upon comparisons of total aspect scores, (input, mediating process, output), within phases of operation for each item of equipment. This increases the number of comparisons possible, and thus yields a more sensitive estimate of reliability.

Per Cent Agreement of Ratings

To further assess the reliability of the Index, the per cent agreement among evaluators at the most microscopic level possible was determined. That is, the evaluation consisted in identifying components, selecting relevant parameters, and determining appropriate dimensions of the parameters. If the same dimensions were always chosen, the scores would be identical, since the data is related to dimensions. It was at the dimensions level that evaluators were compared. First, the total number of dimensions used in an evaluation was determined. Then, the number of times all three evaluators chose the same dimension was determined. The comparison of these two numbers was expressed as "per cent agreement." This information is presented below in Table VII for each equipment and across all equipments.

Table VII. Per Cent Total Agreement Among Three Evaluators

Equipment	Total Entries	Total Number of Agreements	Percentage Agreement
AN/APS-94	427	379	89%
M-33	1023	894	87%
AN/MLQ-8	425	271	64%
All	1875	1544	82%

Validity

The construct validity of the Index, i.e., its measurement of factors critical to operator performance, seems assured. Index scores are a function of factors characterized by an experimentally demonstrated relationship to performance time and operator error in system operation. Content validity, or the extent to which the content of the Index samples factors related to operational complexity, cannot be so easily demonstrated. The literature survey leading to the identification of relevant components, parameters, and dimensions in the Data Store was both systematic and comprehensive. However, it cannot be established that all of the critical factors relevant to operation of electronic equipment have been considered in the experimental literature.

The critical issue is, of course, how well predictions of operator performance based on the Index conform to actual performance in an operational situation. Due to practical limitations involved in the conduct of the study, it was not possible to obtain statistical measures of the predictive validity of the Index. Attempts were made by both the research staff and U. S. Army Electronic Proving Ground personnel to obtain data on actual operator performance with the equipments evaluated. These attempts were hampered by the relative unavailability of operators and their equipment, the time and technical difficulties involved in obtaining microscopic time measures, and the limited number of qualified observers.

In spite of these difficulties, a few performance times were obtained, which were compared with predicted performance times generated through application of the Index. This comparison, though too limited in scope for statistical analysis, did provide some interesting information regarding the extent to which the Operability Index can predict performance <u>time</u>.

First, predicted times seem to be much more accurate for behaviors involving the use of controls and displays than for more gross manipulations, such as cable connections or antenna erection. Predicted times for control panel operations were almost always in close agreement with actual performance times. Observed times for gross manipulations, however, often were as much as three times larger than predicted times. Informal observation during the tryout did suggest that there is tremendous variability in performance time associated with gross behaviors.

Another interesting indication was that observed performance time was almost always greater than predicted time. This trend seemed constant even in control panel operations with experienced personnel, even though the differences between predicted and observed time was small.

It was not possible to obtain formal data concerning operator error during the tryout. The actual observation of operator error was impractical within the scope of the study due to the infrequency of errors. Following the tryout, however, an attempt was made to assess the opinion of experienced operators and their supervisors with regard to the relative potential error associated with each of the four evaluated equipments. Forms and instructions for the ranking of each of the four equipment items were

prepared by the staff and taken into the field by site personnel. Unfortunately, only one operator could be located who had experience with more than one equipment, and this individual was familiar with only two systems.

It was possible to obtain a gross, rational estimate of the general operational complexity of each of the evaluated equipments from the users. Users of the equipment expressed some opinion as to the difficulty of their system as compared with other units which they had operated previously. These users were primarily technical supervisory personnel with several years experience in operating the equipment under a wide range of conditons. It is felt that the qualifications of these personnel, their unanimity of opinion, and the gross design differences between the evaluated equipments provide ample justification for the ranking of these equipments according to their operational complexity. As can be seen from Table VIII, the total time and reliability scores from the Index agree well with this ranking.

Table VIII. Indication of Index Validity

Equipment (Ranked in order of judged complexity	Mean Index Scores	
	Time in Seconds	Reliability
AN/GRC-50 (Judged most complex)	1098.51	.46
M-33	288.48	.81
AN/MLQ-8	181.04	.96
AN/APS-94	104.14	.96

Thus, although a formal, statistical estimate of Index validity was not possible within the scope of the study, the available information suggests that the Index does provide a reasonably valid appraisal of the operating complexity of electronic equipment. Adequate validation, however, yet remains to be done.

Utility

The Sample Equipment Evaluations Report, devoted to reports on the evaluations of all four equipments, demonstrates the utility of the Index. The major uses of the evaluation results are briefly described below.

Acceptability of the level of complexity of equipment is reflected by the total time and reliability estimates provided for the over-all operating sequence. These estimates provide information for deciding upon the acceptability of the given human engineering design of evaluated equipment. The decisions anticipated will be two-fold: 1) can required operations usually be performed with the time expected to be available, and 2) is operator reliability sufficient for the intended mission of the equipment? The necessity of these decisions should be emphasized. The results of the Index are not end processes; they must be interpreted in the light of the total equipment context and the mission the equipment is to achieve. Low reliability and high time scores are neither good nor bad, in and of themselves. To strive for .99 reliability for all equipments and missions is senseless. The reasonable approach is to strive for enough reliability to meet the purpose of the equipment, and no more. The Index cannot make the decisions. It can only facilitate the decisions by providing information of direct relevance.

Redesign alternatives are reflected by the acceptability of the existing design, and the potential for enhancing the acceptability by human engineering redesign. Assuming that redesign is to be seriously considered the Index scores are sufficiently diagnostic that recommendations can be made in detail, and quantitatively justified in terms of enhanced acceptability of the equipment in general.

Selection and training of operators may be either an alternative to redesign, or may be a separate consideration. Selection and training recommendations never fully compensate for design inadequacies that are, in part responsible for the complexity. However, whether or not redesign is a consideration, Index results can provide information relevant to the selection and training of operators so that actual operating performance will be better than the Index estimates indicate.

RECOMMENDATIONS

Versatility

It is unfortunate that the Index could not be tried out on equipment still in the developmental cycle. It is at this stage where Index results would be most beneficial, and, in fact, is the stage at which evaluation with the Index was intended. It was also hoped that a broader sample of equipment could be evaluated. Plans were made toward this end, but the equipment was not available at the time of the tryout. Although current evidence seems to assure the versatility of the Index, it is recommended that further tryouts be conducted on a variety of equipments, and at various stages of the developmental cycle.

Reliability

The current evidence clearly suggests that the Index is a reliable evaluation tool. However, the number of equipments and evaluators available was too limited to insure that the reliability figures presented are accurate. Also, the figures reported are restricted to inter-evaluator reliability. It would be highly desirable to follow this up with test-retest reliability estimates for a number of equipments using an adequate number of evaluators.

Validity

It is in this area that the current study is most restricted. Although the evidence is limited, it would seem that the evaluation procedure is valid, and that it orders equipment in terms of complexity in agreement with expert judges. The actual validity of results, the accuracy of the time and reliability scores, is not established. Some differences in Index scores and actual measures were obtained. The information was insufficient to determine whether this was a consistent or sporadic difference. Consistent differences, if they exist, could be easily remedied by scaling the scores to compensate for the differences. Sporadic differences, unless they relate to inherently variable tasks, could probably be eliminated by altering appropriate instructions. Neither approach is called for on the basis of available information.

Rigorously assessing the validity of the Index results against adequate measures of field operation is of primary importance to the use and further development of the Index. Lacking this validation, the Index will remain a tentative procedure, even though based on the best _a priori_ information currently available.

Utility

While the utility of the Index seems assured, it was not possible to investigate this characteristic with the various personnel and agencies who will make use of Index results. The information provided by the Index seems to be uniquely useful for a variety of purposes. Whether this information would, in fact, be useful to testing, training, and personnel specialists remains in question. It is conceivable that the information provided is of the wrong sort, at the wrong level, or expressed in the wrong terms to be maximally useful. It is equally likely that there are untapped sources of information in the Index. Clearly, the utility of the Index should be determined in association with the people who have need for the information provided.

BIBLIOGRAPHY

Altman, J. W., Marchese, Angeline C., & Marchiando, Barbara W. Guide to design of mechanical equipment for maintainability. Wright-Patterson AFB, Ohio: Aeronautical Systems Division, August 1961. (ASD TR 61-381)

Baker, C. A., & Grether, W. F. Visual presentation of information. Wright-Patterson AFB, Ohio: Aero Medical Laboratory, August 1954. (WADC TR 54-160)

Berkun, M. M., & Van Cott, H. P. Checklist of human engineering evaluation factors. (plans inspection) "CHEEF I." Pittsburgh: American Institute for Research, September 1956.

Craig, R. C., & Purifoy, G. R. Jr. Training for emergency performance. III. Preliminary field study of selected recommendations. Pittsburgh: American Institute for Research, March 1957.

Ely, J. H., Thomson, R. M., & Orlansky, J. Design of controls. Wright-Patterson AFB, Ohio: Wright Air Development Center, November 1956. (WADC TR 56-172)

Fitzpatrick, R. A checklist for human factors in operational suitability tests. Kirtland AFB, N. M.: Air Force Special Weapons Center, August 1955.

Folley, J. D., Jr., & Altman, J. W. Guide to design of electronic equipment for maintainability. Wright-Patterson AFB, Ohio: Wright Air Development Center, April 1956. (WADC TR 56-218)

Kaufman, R. A., Oehrlein, T. F., & Kaufman, M. L. Predicting human reliability--implications for operations and maintenance in space. Paper read at National IAS ARS Joint Meeting, Los Angeles, June 1961.

Krumm, R. L., & Kirchner, W. K. Human factors checklists for test equipment, visual displays, and ground support equipment. Kirtland AFB, N. M.: American Institute for Research, February 1956. (AFSWC TN 56-12)

Miller, R. B., Craig, R. C., & Purifoy, G. R., Jr. Analysis of perceptual interpretive, and judgmental demands of flight personnel during aircraft emergencies. Pittsburgh: American Irstitute for Research, April 1957.

Munger, M. R., & Willis, M. P. Development of an index of electronic maintainability. Pittsburgh: American Institute for Research, 1959.

Shapero, A., & Bates, C., Jr. A method for performing human engineering analysis of weapon systems. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1959. (WADC TR 59-784)

Siegel, A. I., & Wolf, J. J. Techniques for evaluating operator loading in man-machine systems: A description of a model and the results of its first application. Wayne, Pennsylvania: Applied Psychological Services, February 1959.

Van Cott, H. P. Checklist of human engineering evaluation factors (Design inspection) "CHEEF 2." Pittsburgh: American Institute for Research, September 1956.

Van Cott, H. P., & Altman, J. W. Guide to human engineering in the design of weapon systems. Pittsburgh: American Institute for Research, March 1956.

Williams, H. L. Reliability evaluation of the human component in man-machine systems. Falls Church, Virginia: Melpar, Inc., April 1958.

APPENDIX A

DATA STORE REFERENCES

Andreas, B. G., Gerall, A. A., Green, R. F., & Murphy, D. P. Performance in following tracking as a function of the sensitivity of the airplane-type control stick. J. Psychol., 1957, 43, 169-179.

Archer, E. J. Identification of visual patterns as a function of information load. J. Exp. Psychol., 1954, 48, 313-317.

Austin, T. R., & Sleight, R. B. Factors related to speed and accuracy of tactual discrimination. J. Exp. Psychol., 1952, 44, 283-287.

Bahrick, H. P., Bennett, W. F., & Fitts, P. M. Accuracy of positioning responses as a function of spring loading in a control. J. Exp. Psychol., 1955, 49, 437-444.

Bahrick, H. P., Fitts, P. M., & Schneider, R. Reproduction of simple movements as a function of factors influencing proprioceptive feedback. J. Exp. Psychol., 1955, 49, 445-454.

Baker, K. E., & Feldman, H. Threshold-luminance for recognition in relation to frequency of prior exposure. Amer. J. Psychol., 1956, 69, 278-280.

Bald, L., Berrien, F. K., Price, J. B., & Sprague, R. O. Errors in perceiving the temporal order of auditory and visual stimuli. J. Appl. Psychol., 1942, 26, 382-388.

Barber, J. L., & Garner, W. R. The effect of scale numbering on scale reading accuracy and speed. J. Exp. Psychol., 1951, 41, 298-309.

Bartlett, D. J., & Gronow, D. G. C. Manual dexterity and tactile sensitivity in the cold. London: R.A.F., Institute of Aviation Medicine, November 1952. (FPRC-806)

Bartlett, D. J., & Gronow, D. G. C. Effects of heat stress on mental performance. London: R.A.F., Institute of Aviation Medicine, August 1953. (FPRC-846)

Bartlett, N. R., & Williams, S. B. Signal mark size and visibility of radar signals on a plan position indicator. Jamestown, R.I.: Johns Hopkins University, Systems Research Field Laboratory, September 1947. (Prepared for Special Devices Center) (ATI 122-649)

Bauer, H. J. Discrimination of tactual stimuli. J. Exp. Psychol., 1952, 44, 455-459.

Baxter, B. A. A study of reaction time using factorial design. J. Exp. Psychol., 1942, 31, 430-437.

Baxter, B. A., & Travis, R. C. The reaction time to vestibular stimuli. J. Exp. Psychol., 1938, 22, 277-282.

Bradley, J. V. Desirable control-display relationships for moving-scale instruments. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1954. (WADC-TR 54-423)

Bradley, J. V. Effect of knob arrangement on consumption of panel space. Wright-Patterson AFB, Ohio: Wright Air Development Center, June 1956. (WADC-TR 56-202)

Bradley, J. V. Direction of knob turn stereotypes. J. Appl. Psychol., 1959, 43, 21-24.

Bradley, J. V. Tactual coding of cylindrical knobs. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1959. (WADC-TR 59-182)

Bradley, J. V., & Arginteanu, J. Optimum knob diameter. Wright-Patterson AF Ohio: Wright Air Development Center, November 1956. (WADC-TR 56-96)

Bradley, J. V., & Stump, N. E. Minimum allowable dimensions for controls mounted on concentric shafts. Wright-Patterson AFB, Ohio: Wright Air Development Center, December 1955. (WADC-TR 55-355)

Bradley, J. V., & Stump, N. E. Minimum allowable knob crowding. Wright-Patterson AFB, Ohio: Wright Air Development Center, December 1955. (WADC-TR 55-455)

Bradley, J. V., & Wallis, R. A. Spacing of on-off controls. I: Pushbuttons. Wright-Patterson AFB, Ohio: Wright Air Development Center, April 1958. (WADC-TR 58-2)

Bradley, J. V., & Wallis, R. A. Spacing of on-off controls. II: Toggle switches. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1959. (WADC-TR 58-475)

Brown, J. S., Knauft, E. B., & Rosenbaum, G. The accuracy of positioning reactions as a function of their direction and extent. Amer. J. Psychol., 1948, 61, 167-182.

Brown, J. S., & Slater-Hammel, A. T. The effect of speed-up instructions upon the performance of discrete movements in the horizontal plane. Iowa City, Ia: State University of Iowa, January 1948. (Prepared for Special Devices Center) (ATI 26355)

Brown, J. S., Slater-Hammel, A. T., & Bilodeau, E. A. Characteristics of discrete movements in the horizontal plane when executed with one and with two hands. Iowa City, Ia: State University of Iowa, August 1948. (Prepared for Special Devices Center) (ATI 35622)

Caldwell, L. S., & Herbert, M. J. The judgment of angular positions in the horizontal plane on the basis of kinesthetic cues. Fort Knox: Army Medica Research Laboratory, October 1955. (RN-216)

Carr, W. J., & Garner, W. R. The maximum precision of reading fine scales. J. Psychol., 1952, 34, 85-94.

Carter, L. F. The relative effectiveness of presenting numerical data by use of tables and graphs. Army Air Forces Hdqs., Air Materiel Command, Aero Medical Laboratory, April 1946. (AMC TSEAA 694-1)

Carter, L. F. A study of the best design of tables and graphs used for presenting numerical data. Army Air Forces Hdqs., Air Materiel Command, Aero Medical Laboratory, September 1946. (TSEAA-1D)

Chalmers, E. L., Goldstein, M., & Kappauf, W. E. The effect of illumination on dial reading. Wright-Patterson AFB, Ohio: Air Materiel Command, August 1950. (AF TR 6021)

Chapanis, A. Accuracy of oral transmission of target information. Jamestown, R.I.: Johns Hopkins University, Systems Research Field Laboratory, December 1946. (ATI 122 648) (Prepared for Special Devices Center)

Chapanis, A. Speed of reading target information from a direct-reading, counter-type indicator vs. conventional radar bearing-and-range dials. Jamestown, R.I.: Johns Hopkins University, Systems Research Field Laboratory, November 1946.

Chapanis, A. Studies of manual rotary positioning movements: I. The precision of setting an indicator knob to various angular positions. J. Psychol., 1951, 31, 51-64.

Chernikov, R., Duey, J. W., & Taylor, F. V. Two dimensional tracking with identical and different control dynamics in each coordinate. Washington, D. C.: Naval Research Laboratory, November 1959. (NRL R-5424)

Christensen, J. M. The effect of the staircase scale on dial reading accuracy. Wright-Patterson AFB, Ohio: Air Materiel Command, October 1948. (MCREXD 694-1-P)

Christensen, J. M. Quantitative instrument reading as a function of dial design, exposure time, preparatory fixation, and practice. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1952. (WADC-TR 52-116)

Christensen, J. M. The importance of certain dial design variables in quantitative instrument reading. Wright-Patterson AFB, Ohio: Wright Air Development Center, October 1955. (WADC-TR 55-376)

Churchill, A. V. The effect of scale interval length and pointer clearance on speed and accuracy of interpolation. J. Appl. Psychol., 1956, 40, 358-361.

Cohen, J., & Dinnerstein, A. J. A comparison of a linear scale and three logarithmic scales on the time for check reading. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1958. (WADC-TR 57-63)

Cohen, J., & Senders, Virginia L. The effects of absolute and conditional probability distributions on instrument reading, III. A comparison of a linear scale and two scales with expanded central portions. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1958. (WADC-TR 57-65)

Cohen, J., Vanderplas, J. M., & White, W. J. Effect of viewing angle and parallax upon accuracy of reading quantitative scales. J. Appl. Psychol., 1953, 37, 482-488.

Connell, Shirley C. Some variables affecting instrument check reading. Wright-Patterson AFB, Ohio: Air Materiel Command, August 1950. (AF TR 6024)

Conrad, R. Some effects of performance of changes in perceptual load. J. Exp. Psychol., 1955, 49, 313-322.

Crannell, C. W., & Debons, A. Illumination and tilt as factors in the legibility of reflex-reflective numerals. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1958. (WADC-TR 58-47)

Davidson, A. L., Devoe, D. B., Spragg, S. D. S., & Green, R. F. Accuracy of knob setting as a function of: I. The plane in which the knob turns; and II. The diameter of the knob. Special Devices Center, Office of Naval Research, January 1953. (GR SDC 241-6-8)

Davis, L. E. Human factors in design of manual machine controls. Mechanical Engineering, October 1949, 71, 811-816, 837.

Davis, R., Wehrkamp, R., & Smith, K. U. Dimensional analysis of motion: I. Effects of laterality and movement direction. J. Appl. Psychol., 1951, 35, 363-366.

Deese, J. Signal size and detectability on a PPI display. Wright-Patterson AFB, Ohio: Wright Air Development Center, August 1954. (WADC-TR 54-166)

Deese, J., & Ormond, Elizabeth. Studies of detectability during continuous visual search. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1953. (WADC-TR 53-8)

Duncker, K. Problem solving. Psych. Monographs, 1945, 58.

Dunlap and Associates, Inc. Investigation of speed and accuracy in reading multiscale meters. Stamford, Conn.: April 1960.

Elkin, E. H. Effects of scale shape, exposure time, and display-response complexity on scale reading efficiency. Wright-Patterson AFB, Ohio: Wright Air Development Center, February 1959. (WADC-TR 58-472)

Elliott, D. N., & Howard, E. F. Effect of position upon warning light effectiveness. Perception & Motor Skills, June 1956, 69.

Forbes, G. The effect of certain variables on visual and auditory reaction time. J. Exp. Psychol., 1945, 35, 153-162.

Ford, A. Types of errors in location judgments on scaled surfaces. I. Errors of configuration. J. Appl. Psychol., 1949, 33, 373-381.

Gardner, J. F. Direction of pointer motion in relation to movement of flight controls - cross-pointer type instrument. Wright-Patterson AFB, Ohio: Air Materiel Command, June 1950. (AF TR 6016)

Garvey, W. D. The effects of "task-induced stress" on man-machine systems performance. Washington, D. C.: Naval Research Laboratory, September 1957. (R-5015)

Garvey, W. D., & Knowles, W. B. Response time patterns associated with various display-control relationships. J. Exp. Psychol., 1954, 47, 315-322.

Garvey, W. D., & Mitnick, L. L. Effect of additional spatial references on display-control efficiency. J. Exp. Psychol., 1955, 50, 276-282.

Gerall, A. A., Sampson, P. B., & Spragg, S. D. S. Performance on a tracking task as a function of position, radius, and loading of control cranks: I. Stationery targets. J. Psychol., 1956, 41, 135-150.

Graham, Norah E. The speed and accuracy of reading horizontal, vertical, and circular scales. J. Appl. Psychol., 1956, 40, 228-232.

Green, R. F., Norris, E. B., & Spragg, S. D. S. The effect on performance of varying the directions and planes of movement of the control cranks on a compensatory tracking task. Rochester, N. Y.: University of Rochester, August 1955. (SDC TR 241-6-18) (Prepared for Special Devices Center)

Gregg, L. W. The effect of stimulus complexity on discrimination responses. J. Exp. Psychol., 1954, 48, 289-297.

Grether, W. F. Speed and accuracy of dial reading as a function of dial diameter and spacing of scale divisions. Wright-Patterson AFB, Ohio: Air Materiel Command, March 1947. (TSEAA 694-1E)

Grether, W. F. The effect of variations in indicator design upon speed and accuracy of altitude readings. Army Air Forces Hdqs., Air Materiel Command, Aero Medical Laboratory, September, 1947. (TSEAA 694-14)

Grether, W. F. Psychological factors in check reading of single instruments. Wright-Patterson AFB, Ohio: Air Materiel Command, September 1949. (MCREXD 694-17A)

Grether W. F. Psychological factors in instrument reading. I. The design of long-scale indicators for speed and accuracy of quantitative readings. J. Appl. Psychol., 1949, 33, 363-372.

Grether, W. F., & Williams, A. C., Jr. Psychological factors in instrument reading. II. The accuracy of pointer position interpolation as a function of the distance between scale marks and illumination. J. Appl. Psychol., 1949, 33, 594-604.

Harriman, M. W. Visibility of cathode-ray tube screens: Search time as a function of signal strength. J. Psychol., 1950, 29, 247-250.

Hartman, B. O. The effect of joystick length on pursuit tracking. Fort Knox: Army Medical Research Laboratory, November 1956. (R-279)

Hartman, B. O. The effect of the extent of movement (control sensitivity) on pursuit tracking performance. Fort Knox: Army Medical Research Laboratory, May 1957. (R-308)

Hartman, B. O., Jaynes, W. E., & Herbert, M. J. Analysis of abductive and adductive phases of movement in continuous tracking. Fort Knox: Army Medical Research Laboratory, December 1957. (R-314)

Hartman, B. O., & McKenzie, R. E. An exploratory study of the effects of low work rates on performance during overload. Brooks AFB, Texas: School of Aviation Medicine, Aerospace Medical Center. (unpublished)

Hartman, B. O., & McKenzie, R. E. Systems operator proficiency, effects of speed stress on overload performance. Brooks AFB, Texas: School of Aerospace Medicine, Aerospace Medical Center, June 1961. (GR AF SAM 61-40)

Henneman, R. H., Lewis, Penelope, & Matthews, T. L. Influence of the sensory requirements of the distracting task. Wright-Patterson AFB, Ohio: Wright Air Development Center, November 1953. (WADC-TR 53-309)

Henneman, R. H., & Matthews, T. L. The influence of message length and distracting task complexity. Wright-Patterson AFB, Ohio: Wright Air Development Center, April 1955. (WADC-TR 54-145)

Herbert, M. J. The speed and accuracy with which six linear arm movements can be visually positioned from two different control locations. Fort Knox: Army Medical Research Laboratory, March 1957. (R-260)

Holland, J. G., & Lee, W. The influence of message distortion and message familiarity. Wright-Patterson AFB, Ohio: Wright Air Development Center, April 1955. (WADC-TR 54-287)

Horton, G. P. Accuracy of reading target location and size of schematic PPI display. Wright-Patterson AFB, Ohio: Air Materiel Command, October 1949. (AF TR 5961)

Horton, G. P. An analysis of errors made on a schematic PPI display. Wright-Patterson AFB, Ohio: Air Materiel Command, October 1949. (AF TR 5960)

Howland, D. An investigation of the performance of the human monitor. Wright-Patterson AFB, Ohio: Wright Air Development Center, July 1958. (WADC-TR 57-431)

Howland, D., & Noble, M. E. The effect of physical constants of a control on tracking performance. J. Exp. Psychol., 1953, 46, 353-360.

Humphries, M. Performance as a function of control-display relations, positions of the operator, and locations of the control. J. Appl. Psychol., 1958, 42, 311-316.

Hunt, D. P. Tracking performance as a function of feedback specificity. Wright-Patterson AFB, Ohio: Aero Medical Laboratory, March 1959. (WADC-TR 58-584)

Hunt, D. P., & Warrick, M. J. Accuracy of blind positioning of a rotary control. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1957. (WCLD-TN 52-106)

Jenkins, W. L. Mean least turn and its relation to making settings on a linear scale. Wright-Patterson AFB, Ohio: Wright Air Development Center, May 1957. (WADC-TR 57-210)

Jenkins, W. L., & Connor, M. B. Some design factors in making settings on a linear scale. J. Appl. Psychol., 1949, 33, 395-409.

Jenkins, W. L., & Karr, A. C. The use of a joystick in making settings on a simulated scope face. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1954. (WADC-TR 53-430)

Jenkins, W. L., & Karr, A. C. The influence of viewing distance in making settings on a linear scale. Wright-Patterson AFB, Ohio: Wright Air Development Center, November 1955. (WADC-TN 55-204)

Jenkins, W. L., & Olson, M. W. The use of levers in making settings on a linear scale. Wright-Patterson AFB, Ohio: Wright Air Development Center, August 1951. (WADC-TR-6563)

Jenkins, W. O. The accuracy of pilots in applying pressures on rudder pedals. Army Air Forces Hdqs., Air Materiel Command, Aero Medical Laboratory, September 1946. (TSEAA 694-3B)

Jenkins, W. O. The accuracy of pilots and non-pilots in applying pressures on a control stick. Wright-Patterson AFB, Ohio: Air Materiel Command, August 1946. (ATI 105-858)

Jenkins, W. O. The discrimination and reproduction of motor adjustments with various types of aircraft controls. Amer. J. Psychol., 1947, 60, 397-406.

Jerison, H. J. Duration of vigil and the decrement function. Wright-Patterson AFB, Ohio: Wright Air Development Center, December 1958. (WADC-TR 58-369)

Jerison, H. J., & Wallis, R. A. One clock and three clock monitoring. Wright-Patterson AFB, Ohio: Wright Air Development Center, April 1957. (WADC-TR 57-206)

Jerison, H. J., & Wallis, R. A. Performance on a simple vigilance task in noise and in quiet. Wright-Patterson AFB, Ohio: Wright Air Development Center, June 1957. (WADC-TR 57-318)

Jerison, H. J., & Wing, S. Effects of noise and fatigue on a complex vigilance task. Wright-Patterson AFB, Ohio: Wright Air Development Center, January 1957. (WADC-TR 57-14)

Kappauf, W. E. Design of instrument dials for maximum legibility: Part V. Origin location, scale break, number location, and contrast direction. Wright-Patterson AFB, Ohio: Wright Air Development Center, May 1951. (AF TR 6366)

Kappauf, W. E., & Smith, W. A. Design of instrument dials for maximum legibility: Part II. A preliminary experiment on dial size and graduation. Wright-Patterson AFB, Ohio: Aero Medical Laboratory, July 1948. (MCREXD 694-1-N)

Kappauf, W. E., & Smith, W. A. Design of instrument dials for maximum legibility. Part III. Some data on the difficulty of quantitative reading in different parts of a dial. Wright-Patterson AFB, Ohio: Air Materiel Command, May 1950. (AF TR 5914)

Karpovich, P. V., Dering, R. R., & Ikai, M. Reflex and reaction time. Springfield, Mass.: Springfield College, July 1960. (AD 240597) (Prepared for Army Medical Research & Development Command)

Kaufman, E. L., Reese, T. W., Volkmann, J., & Rogers, S. Accuracy, variability, and speed of adjusting an indicator to a required bearing. Hadley, Mass.: Mt. Holyoke College, September 1947. (166-I-MHC-4) (Prepared for Special Devices Center)

Klemmer, E. T. A further study of information transmission with matrix patterns. Bolling AFB, Washington, D. C.: Air Research & Development Command, September 1957. (AFCRC TN 57-1)

Knowles, W. B., Garvey, W. D., & Newlin, E. P. The effect of speed and load on display-control relationships. J. Exp. Psychol., 1953, 46, 65-75.

Krulee, G. K., & Weisz, A. Studies in the visual discrimination of multiple-unit displays. J. Exp. Psychol., 1955, 50, 316-324.

Kurke, M. I. Evaluation of a display incorporating quantitative and check reading characteristics. J. Appl. Psychol., 1956, 40, 233-236.

Lawrence, D. H., & LaBerge, D. L. Relationship between recognition accuracy and order of reporting stimulus dimensions. J. Exp. Psychol., 1956, 51, 12-1

Lazar, R. G., & Williams, J. R. Investigation of natural movements in azimuth and elevation lever control adjustments for horizontal and vertical positions. Aberdeen Proving Ground, Md.: Army Ordnance Human Engineering Laboratory, April 1959. (APG TM 3-59)

Leyzorek, M. Mounting angle of a VJ remote radar indicator and its effect on operator performance. Jamestown, R.I.: Johns Hopkins University, Systems Research Field Laboratory, February 1948. (Prepared for Special Devices Center)

Lincoln, R. S., & Averbach, E. Spatial factors in check reading of dial groups. J. Appl. Psychol., 1956, 40, 105-108.

McGuire, J. C. The effect of target velocity and the area of error-tolerance circles upon performance in two-dimensional compensatory tracking task. Wright-Patterson AFB, Ohio: Wright Air Development Center, August 1954. (WADC-TR 54-431)

Mackworth, N. H. Researches on the measurement of human performance. London: Medical Research Council, 1950. (Special Report Series No. 268)

Mitchell, M. J. H., & Vince, M. A. The direction of movement of machine controls Quar. J. Exp. Psychol., 1951, 3, 24-35.

Muckler, F. A. Man-machine tracking performance with short-period oscillatory control system transients. Wright-Patterson AFB, Ohio: Wright Air Developme Division, January 1960. (GR AF WADD-TR 60-3)

Newton, J. M. An investigation of tracking performance in the cold with two types of controls. Fort Knox: Army Medical Research Laboratory, September 1957. (R-324)

Noble, M. E., Fitts, P. M., & Marlowe, E. The inter-relations of task variables in continuous pursuit tasks: II. Visual-display scale, arm-control scale, and target frequency in compensatory tracking. Lackland AFB, Texas: Human Resources Research Center, ARDC, December 1953. (AF HRRC RB 53-55)

Office of Scientific Research and Development, National Defense Research Committee, Div. 7. Inertia, friction, and diameter in handwheel tracking. September 1943. (GR OSRD RS 74 R-4)

Peters, R. W. Research on psychological parameters of sound. Wright-Patterson AFB, Ohio: Wright Air Development Division, February 1960. (WADD-TR 60-249)

Pollack, I., & Tecce, J. Speech annunciator warning indicator system: Preliminary evaluation. J. Acoustical Society of America, January 1958, 30, (1), 58-61.

Rappaport, M. The role of redundancy in the discrimination of visual forms. J. Exp. Psychol., 1957, 53, 3-10.

Rockway, M. R. The effect of variations in control/display ratio and exponential time delay on tracking performance. Wright-Patterson AFB, Ohio: Wright Air Development Center, December 1954. (WADC-TR 54-618)

Ross, S., Shepp, B. E., & Andrews, T. G. Response preferences in display-control relationships. J. Appl. Psychol., 1955, 39, 425-428.

Rubin, G., Von Trebra, Patricia, & Smith, K. U. Dimensional analysis of motion: III. Complexity of movement pattern. J. Appl. Psychol., 1952, 36, 272-276.

Russell, R. W. Effects of variations in ambient temperature on certain measures of tracking skill and sensory sensitivity. Fort Knox: Army Medical Research Laboratory, February 1957. (RN 300)

Schapiro, H. B. Factors affecting legibility of digits. Wright-Patterson AFB, Ohio: Wright Air Development Center, June 1952. (WADC-TR 52-127)

Senders, Virginia L. The effect of number of dials on qualitative reading of a multiple dial panel. Wright-Patterson AFB, Ohio: Wright Air Development Center, November 1952. (WADC-TR 52-182)

Senders, Virginia L., & Cohen, J. The effects of absolute and conditional probability distributions of instrument settings on scale reading: Repeated exposures of the same setting. Wright-Patterson AFB, Ohio: Wright Air Development Center, October 1954. (WADC-TR 54-253)

Senders, Virginia L., Cohen, J., & Arginteanu, Mary. The effects of absolute and conditional probability distribution on instrument reading. II: A comparison of a linear and logarithmic scale. Wright-Patterson AFB, Ohio: Wright Air Development Center, July 1955. (WADC-TR 54-253)

Siddall, G. J., Holding, D. H., & Draper, J. Errors of aim and extent in manua point to point movement. Occupational Psychol., July 1957, 31.

Sidorsky, R. Precise control manipulation as a function of control location. Fort Knox: Army Medical Research Laboratory, January 1957. (RN 288)

Sidowski, J. B., Morgan, R., & Eckstrand, G. Influence of task complexity and instructions upon simple and discrimination reaction times. J. Exp. Psychol 1958, 55, 163-166.

Simon, C. W. Instrument control configurations affecting performance in a compensatory pursuit task. Wright-Patterson AFB, Ohio: Wright Air Development Center, February 1952. (AF TR 6015)

Simon, J. R., & Simon, Betty Pearl. Duration of movements in a dial setting task as a function of the precision of manipulation. J. Appl. Psychol., 1959, 43, 389-394.

Sleight, R. B. The effect of instrument dial shape on legibility. J. Appl. Psychol., 1948, 32, 170-188.

Smader, R., & Smith, K. U. Dimensional analysis of motion: VI. The component movements of assembly motions. J. Appl. Psychol., 1953, 37, 308-314.

Solley, C. M., & Snyder, F. W. Information processing and problem solving. J. Exp. Psychol., 1958, 55, 384-387.

Spragg, S. D. S. Dial reading performance as related to illumination variables I. Intensity. Wright-Patterson AFB, Ohio: Air Materiel Command, October 1948. (MCREXD 694-21)

Spragg, S. D. S., & Devoe, D. B. The accuracy of control knob settings as a function of size of angle to be bisected and type of end-point cue. Perception & Motor Skills, 1956, 6, 25-28.

Spragg, S. D. S., & Rock, M. L. Dial reading performance as related to illumination variables: II. Spectral distribution. Wright-Patterson AFB, Ohio: Aero Medical Laboratory, December 1948. (MCREXD 694-21A)

Spragg, S. D. S., & Rock, M. L. Dial reading performance as a function of brightness. J. Appl. Psychol., 1952, 36, 128-137.

Steedman, W. C., & Baker, C. A. Target size and visual recognition. Wright-Patterson AFB, Ohio: Wright Air Development Division, February 1960. (WADD-TR 60-93)

Stump, N. E. Toggle switches - activation time as a function of the plane of orientation and the direction of movement. Wright-Patterson AFB, Ohio: Wright Air Development Center, September 1952. (WADC-TR 52-51)

Stump, N. E. Manipulability of rotary controls as a function of knob diameter and control orientation. Wright-Patterson AFB, Ohio: Wright Air Developmen Center, February 1953. (WCRD TN 53-12)

Taubman, R. E. Studies in judged number: I. The judgment of auditory number. J. Gen. Psychol., 1950, 43, 167-194.

Taubman, R. E. Studies in judged number: II. The judgment of visual number. J. Gen. Psychol., 1950, 43, 195-219.

Teichner, W. H., & Zigler, M. J. A method of studying the tactual-kinesthetic sensitivity of the hand. Natick, Mass.: Quartermaster Research & Development Center, November 1953. (RN 244)

Tinker, M. A. Readability of mathematical tables. J. Appl. Psychol., 1954, 38, 436-442.

Tinker, M. A., & Paterson, D. G. Readability of mixed type forms. J. Appl. Psychol., 1946, 30, 631-637.

U. S. Air Force. Psychological aspects of instrument display: I. Analysis of 270 "pilot-error" experiences in reading and interpreting aircraft instruments. Wright-Patterson AFB, Ohio: Air Materiel Command, October 1947. (TSEAA 694-12A)

Vince, M. A. The intermittency of control movements and the psychological refractory period. Brit. J. Psychol., 1948, 38, 149-157.

Vince, M. A. Learning and retention of an "unexpected" control-display relationship under stress conditions. Cambridge, England: Applied Psychology Research Unit, August 1950. (APU RN 125-50)

Von Trebra, Patricia, & Smith, K. U. Dimensional analysis of motion: IV. Transfer effects and direction of movement. J. Appl. Psychol., 1952, 36, 348-353.

Warrick, M. J. Direction of movement in the use of control knobs to position visual indicators. Wright-Patterson AFB, Ohio: Air Materiel Command, April 1947. (TSEAA 694-4C)

Warrick, M. J. Effects of motion relationships on speed of positioning visual indicators by rotary control knobs. Wright-Patterson AFB, Ohio: Air Materiel Command, July 1949. (AF TR 5812)

Warrick, M. J. Counters for airborne use. Wright-Patterson AFB, Ohio: Wright Air Development Center, June 1954. (WADC-TR 54-266)

Warrick, M. J., & Grether, W. F. The effect of pointer alignment on check reading of engine instrument panels. Wright-Patterson AFB, Ohio: Air Materiel Command, June 1948. (MCREXD 694-17)

Wehrkamp, R., & Smith, K. U. Dimensional analysis of motion: II. Travel distance effects. J. Appl. Psychol., 1952, 36, 201-206.

Weiss, B. The role of proprioceptive feedback in positioning responses. J. Exp. Psychol., 1954, 47, 215-244.

Weiss, B. Movement error, pressure variation, and the range effect. J. Exp. Psychol., 1955, 50, 191-196.

White, W. J. The effect of dial diameter on ocular movements and speed and accuracy of check reading of groups of simulated engine instruments. Wright-Patterson AFB, Ohio: Air Materiel Command, June 1949. (R 5826)

White, W. J. The effect of pointer design and pointer alignment position on the speed and accuracy of reading groups of simulated engine instruments. Wright-Patterson AFB, Ohio: Wright Air Development Center, July 1951. (AF TR 6014)

White, W. J., & Riley, M. B. The effects of positive acceleration on the relation between illumination and instrument reading. Wright-Patterson AFB, Ohio: Wright Air Development Center, November 1958. (WADC-TR 58-332)

White, W. J., & Sauer, Shirley C. Scale design for reading at low brightness. Wright-Patterson AFB, Ohio: Wright Air Development Center, March 1954. (WADC-TR 53-464)

APPENDIX B

RANK ORDER RELIABILITY TABLES

Table I. Phase Level Comparison Over Two Equipments

A. Time Scores

Phases	Rank Order by Evaluators A	B	C	ρ	Median ρ
AN/APS-94	2	1	1	AB = .7714	
Chief Radar Operator	3	2	2		
Computer Operator	1	3	3	AC = .8286	.86
Range Operator	4	5	4		
Azimuth Operator	5	4	5	BC = .9429	
Elevation Operator	6	6	6		

B. Reliability Scores

Phases	A	B	C	ρ	Median ρ
AN/APS-94	5	4.5	4	AB = .9857	
Chief Radar Operator	4	4.5	5		
Computer Operator	6	6	6	AC = .9429	.96
Range Operator	2.5	2.5	2.5		
Azimuth Operator	2.5	2.5	2.5	BC = .9857	
Elevation Operator	1	1	1		

Table II. Aspect and Phase Level Comparison for the M-33

A. Time Scores

Aspect or Phase	Rank Order by Evaluator A	B	C	ρ	Median ρ
Total					
Input	3	2	2		
Mediating Process	1	12	13		
Output	4	1	1		
Chief Radar Operator				AB = .5470	
Input	5	3	3		
Mediating Process	15.5	15	15	AC = .6314	.72
Output	7	5	4		
				BC = .8927	
Azimuth Operator					
Input	9	7	8		
Mediating Process	15.5	16	16		
Output	10	8	9		
Range Operator					
Input	11	10	11		
Mediating Process	17	17.5	17.5		
Output	8	6	6		
Elevation Operator					
Input	14	3	12		
Mediating Process	18	17.5	17.5		
Output	13	9	10		
Computer Operator					
Input	12	11	7		
Mediating Process	2	14	14		
Output	6	4	5		

Table II. Continued

B. Reliability Scores

Aspect or Phase	Rank Order by Evaluator A	B	C	ρ	Median ρ
Total					
Input	18	18	18		
Mediating Process	13.5	5.5	5.5	AB = .8457	
Output	17	17	17		
Chief Radar Operator				AC = .8457	.92
Input	6.5	13	13		
Mediating Process	6.5	5.5	5.5	BC = 1.0000	
Output	13.5	15	15		
Azimuth Operator					
Input	6.5	5.5	5.5		
Mediating Process	6.5	5.5	5.5		
Output	6.5	5.5	5.5		
Range Operator					
Input	6.5	5.5	5.5		
Mediating Process	6.5	5.5	5.5		
Output	6.5	13	13		
Elevation Operator					
Input	6.5	5.5	5.5		
Mediating Process	6.5	5.5	5.5		
Output	6.5	5.5	5.5		
Computer Operator					
Input	16	16	16		
Mediating Process	6.5	5.5	5.5		
Output	15	13	13		

Table III. Aspect and Phase Level Comparison for the AN/MLQ-8 (XL-2)

A. Time Scores

Aspect or Phase	Rank Order by Evaluator A	B	C	ρ	Median ρ
Total					
Input	4.5	4.5	6.5	AB = .9762	
Mediating Process	6	6	4		
Output	1	1	1	AC = .7381	.86
Erect Antenna				BC = .8095	
Input	omit	omit	omit		
Mediating Process	8	7	5		
Output	2	2	2		
Operate					
Input	4.5	4.5	6.5		
Mediating Process	7	8	8		
Output	3	3	3		

B. Reliability Scores

Aspect or Phase	A	B	C	ρ	Median ρ
Total					
Input	3.5	3	3		
Mediating Process	3.5	3	3	AB = .8750	
Output	8	8	8	AC = .8750	.94
Erect Antenna					
Input	omit	omit	omit	BC = 1.0000	
Mediating Process	3.5	3	3		
Output	7	6.5	6.5		
Operate					
Input	3.5	3	3		
Mediating Process	3.5	3	3		
Output	3.5	6.5	6.5		

Table IV. Aspect and Phase Level Comparison for the AN/APS-94

A. Time Scores

Phase	Rank Order by Evaluator			ρ	Median ρ
	A	B	C		
Total only					
Input	2	1	2	AB = .5000	
Mediating Process	3	3	3	AC = 1.0000	.7500
Output	1	2	1	BC = .5000	

B. Reliability Scores

Phase	A	B	C	ρ	Median ρ
Total only					
Input	1.5	2	1.5	AB = .8750	
Mediating Process	1.5	1	1.5	AC = .10000	.9375
Output	3	3	3	BC = .8750	